

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: SmithKline Beecham plc et al
- (ii) TITLE OF THE INVENTION: Novel compounds
- (iii) NUMBER OF SEQUENCES: 19
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: SmithKline Beecham
 - (B) STREET: Two, New Horizons Court, Great West Road
 - (C) CITY: Brentford
 - (D) STATE:
 - (E) COUNTRY: UK
 - (F) ZIP: TW8 9EP
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Valentine, Jill B
 - (B) REGISTRATION NUMBER:
 - (C) REFERENCE/DOCKET NUMBER: P31731

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 0181-9752000
- (B) TELEFAX: 0181-9756294
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7193 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

| | |
|---|------|
| CCATGGCGGG CGGCGGCTGC CCCGAGCCT CGGCCGACC GGTGACCAGG ACCACCCCGG | 60 |
| TGGGATAGTG GCCCGCCACC CGGCGCAGCA GACTCCCGGA CACGGACCCG TGGGTGTGCG | 120 |
| CGGAAAGGCC CGGAGGCCGG GTCACAGCCA CGGTAACGC GCGGTGTCCT TGCCCGCGTA | 180 |
| ATCGGGGTCC AGATAGACGA AGGCCCGGTG GACGAGGAAG TCCCGCACCT CGTAGACCGT | 240 |
| GCACCAGCGC CCGGCGGCCC ACTCGGGGTC ACCCGCCCGC CACGGCCCGT CCCGGTGCTC | 300 |
| ACCGTGGGTG GTGCCCTCCG CGGCGAGGAG TTCGGTCCCG GTCAGAATCC AGTTGACGGA | 360 |
| CCACAGATGG TGGGTGATCG AGCGGATGGT GCCCCCGAGG TCGTCGAAGA GCCGGGCGAT | 420 |
| CTCGGACTTG CCCCCGGCCA GACCCCACTT GGGGAAGAAG AAGACCGCGT CCTCGGCGAA | 480 |
| GTAGTCGATC GCGGGGGTGC CGTCGCTGCC GACGCCGCCG TTGTGCAACG CCTTGAAGTA | 540 |
| CGCGGTGATG ACCGCCTTGC GCTGCTCGTC CGTCATACCG GCCGATGCCA CGGACATGAA | 600 |
| ACGACCTCCA GAGATTCCGG GTGGCTGTGC TGGGGCTGCG GAAGGGGTGT CCCCCGCGAA | 660 |
| GGACGGCGGA CGCCGCGGAC GCCGCGGCCG TCTCCCCGGC GGACGGGTCC CAGCGTCCTG | 720 |
| GAGAGGGCTT GCGGGCGGCT TGACGCCGTG CTGTCCCGCG GCTTGCGGAA CGCGAAGTAC | 780 |
| CGGCCAGCGT ACGGGCGTTG CACCGGACGT GTACGCCGGT CGGGACCCCT CGTACCCCCG | 840 |
| GAGCCGGCCG ACCCGGCGG CTCCGGGGGT ACGGACGCGC CGGACCGGCC CGAGCGAGCC | 900 |
| GGACGGGTG GACGGTGCGC GTGGTTCCGG TGTGTCGGAC AGCTCGGACG GACCGGACGG | 960 |
| TGCGCGTGGT TCCGGTGTGT CGGACAGCTC GGACGGGTG GACCGTGCGC GTGGTTCCGG | 1020 |
| CACGCCGAC GGGTCAGTTG CCGATCATGG CGAGCAATGC CGGGGTGTAC CGCTCCCCCG | 1080 |
| ACACCGGGTG GGAGATCGCG GCCGTCACCT CCGCGAGGGA CCGGTCGTCC AGCCGGATCG | 1140 |
| AGGCGGCGGC GAGATTGTCC GCGAGATGGG CCGGGTTTCG GGTGCCCCGG ATCGGGACGA | 1200 |
| CGTCCTCGCC CCGGTGGTGC AGCCAGGCGA GCGCGAGCTG TGCCAGGGTC AGCCCCAGAC | 1260 |

CGTCCGCGAC CGGGCGCAGC CGGTGCAGCA ACGAGCGGTT GCGCGCGAGG GCCGGAGCGC 1320
TGAACCGGGG CTGGCCCCGG CGGAAGTCCT CGTCCCCCAG ATCGTCGGTG GTGCGGATGG 1380
TGCCGGTGAG AAAACCCCGT CCCAGAGGGG CGTAAGCGAC GATCCCGATC CCCAGCTCCC 1440
GGCAGACGGG CACCACCTCG TCCTCGATCC CGCGCGACCA CAGGCTCCAC TCGCTCTGCA 1500
CCGCCGTCAC CGGGTGACCC GCGTCCGCCC GCGCGAGCGT GGCCGCGGAG GGCTCGGAGA 1560
GACCGAGCCT GCGGACCTTG CCTTCGCGCA CCAGCTCGGC CACCGCACCC ACGGTCTCCT 1620
CGATCGGCAC CGCCGGGTCC GTCCAGTGCT GGTAGTACAG GTCGATGCGG TCGGTGCCGA 1680
GACGACGCAG GGACCGTTTC CAGGCCGCGC GGACGTAGGA CGGCTCGCCG CACAAGCCCT 1740
GGGAGGCGCC GTCGGACGAG CGCACCATGC CGAACTTGGT GCGGATCAGC ACCTCGTCCC 1800
GGCGGCCCGC GACCGCCCGT CCGAGCAGCT CCTCACCAGC GCCGAGCCCC TGGACGTCGG 1860
CGGTGTCCAG CAGGGTGACC CCGGCGTCGA CCGCGGCGCG GATGGTGGCC GTCGCCCGGG 1920
CGCGGTCCGG GCGTCCGTAG AAGTCGGTGG TCGGCAGGCA GCCGAGCCCC TGGGCACTGA 1980
CCGGAAGGTC CCGCAGGGCG CGGACCGGCG GACGCGGAAC CGCGGCGGAC ACGGAACCGG 2040
CCGGGGACTC GGGCGGAGAG CGGGACATAC GGAACCTCCA CAGGCGGAGC CGGGAACGGG 2100
ACGAGGGCGA GGACGGGACG GAACGAAGGA GAGGACGGGA CGGACAGCAC GGACGGGACG 2160
GACGGAACGG AGTCGGGAAC CGGGGGGGGT GACCGGAACC GGGCCGTCTT TGGCCCTCCC 2220
CCGTCTCTCC CGCCATCCGC CGTTCTCCCC CGTTCCCTCT CCGTCTCTCC AGCCAACACC 2280
GCCGCCCTTT CCAAGCGCTT GACACGGCAC CGACAGCCGC CGCCGGGCGC CCGATGGGGA 2340
CCCGTGCCCG CCGGTGAGCG GCGGTGAGCG CCGGTACGGG ACCCCACGCG CCGCCGCCCG 2400
GGCGCCCGCC AGGGCCCGCG CGGCCACCCC GGCCCGCCCC GGCCCGAGCG GCGATCCGGG 2460
CCGCTCGCTG CAAGAGGAAC ATCCACAGCC GCACAAGGAG CGCTCCGCAC AGTGGGCACC 2520
ACGTCCGCCC CGTCCCCCACC ACCGTGGCCG GTCCCCACCG GACAGCACAG CACCGCACAG 2580
CACCACATCG CACGGCACAG CACAGCACCA CCGGCACGAG GAACCAAGGA AAGGAACCAC 2640
ACCACCATGA CCTCAGTGGA CTGCACCGCG TACGGCCCCG AGCTGCGCGC GCTCGCCGCC 2700
CGGCTGCCCC GGACCCCCCG GGCCGACCTG TACGCCTTCC TGGACGCCGC GCACACAGCC 2760
GCCGCCTCGC TCCCGGCGC CCTCGCCACC GCGCTGGACA CCTTCAACGC CGAGGGCAGC 2820
GAGGACGGCC ATCTGCTGCT GCGCGGCCTC CCGGTGGAGG CCGACGCCGA CCTCCCCACC 2880
ACCCCGAGCA GCACCCCGGC GCCCGAGGAC CGCTCCCTGC TGACCATGGA GGCCATGCTC 2940
GGACTGGTGG GCCCGCGGCT CGGTCTGCAC ACGGGGTACC GGGAGCTGCG CTCGGGCACG 3000
GTCTACCACG ACGTGTAACC GTCGCCCGGC GCGCACCACC TGTCTTCGGA GACCTCCGAG 3060
ACGCTGCTGG AGTTCCACAC GGAGATGGCC TACCACCGGC TCCAGCCGAA CTACGTCATG 3120
CTGGCCTGCT CCCGGGCCGA CCACGAGCGC ACGGCGGCCA CACTCGTCGC CTCGGTCCGC 3180
AAGCGCTGC CCCTGCTGGA CGAGAGGACC CGGGCCCGGC TCCTCGACCG GAGGATGCCC 3240
TGCTGCGTGG ATGTGGCCTT CCGCGGCGGG GTGGACGACC CGGGCGCCAT CGCCAGGTC 3300
AAACCGCTCT ACGGGGACGC GGACGATCCC TTCCTCGGGT ACGACCGCGA GCTGCTGGCG 3360
CCGAGGACC CCGCGGACAA GGAGGCCGTC GCCGCCCTGT CCAAGGCGCT CGACGAGGTC 3420
ACGGAGGCGG TGTATCTGGA GCCCGCGGAT CTGCTGATCG TCGACAACTT CCGCACCACG 3480
CACGCGCGGA CGCCGTTCTC GCCCCGCTGG GACGGGAAGG ACCGCTGGCT GCACCGCGTC 3540
TACATCCGCA CCGACCGCAA TGGACAGCTC TCCGGCGGCG AGCGCGCGGG CGACGTCGTC 3600
GCCTTCACAC CGCGCGGCTG AGTCCCGGG TCCGACACCG CGCGGCTGAA CCCACGGTCC 3660

| | | | | | | |
|------------|-------------|------------|------------|------------|------------|------|
| GGGGCCCACG | GTCCGGCACC | GCGCGGCTGA | GCCCCCGGGT | CCGGCAGCGG | GCGGCTGAAC | 3720 |
| CCCCGCCCCG | GGCCACCGCC | CGACCGCCCC | CGCGCACCGG | ACGCGCCCCG | CTGTACGGCG | 3780 |
| GTCCCGCCCC | GGCCCGTACA | CCTGAAGCGC | CCGGCGGACC | GCCGCCCCCG | CGGGGGACGG | 3840 |
| ACAGAGCCCG | GTGCGGGAGG | ACGTCTCTCC | GCACCCGGCT | CCCACCGTTC | CGCACCGACC | 3900 |
| GCACCCGACC | GTGCCGCAGG | CGCCACCGGC | ACCGCACCGC | CCGCGCCGGC | AGCCACCACA | 3960 |
| GGCGCCACGC | CGCCCGCACG | GTGCCCCGCG | TGCTCAGCCC | CCGTCCACCG | GGCTGTCCAG | 4020 |
| CAGCCGCCCC | AGCGCGCCCC | CGATGAACTC | CCGGTCGGCG | GCCGACCCCC | CGGACCCCGC | 4080 |
| GAGATGCCCC | CACACTCCCC | GGATCACCTC | CAGCGAGGCA | TACGGCAGCA | GATCGGCCAC | 4140 |
| CCGCTTCTCG | TCCTCGACGG | CGAAACACAC | GTCCAGGGCG | CCCGGCAGCA | CCACGGCCCC | 4200 |
| CGCCGTGACG | GAGGCCAGCG | CCGCCTCGAC | GCTCCCCCGG | GCCCCGGGTG | TCGCCCCCAC | 4260 |
| ATCCGTGTTC | TCCCAGGTGC | GCACCATGGT | GAGCAGATCC | GCGGCGCCGG | GCCCCGAGAG | 4320 |
| GAAGACCTGC | TCCCAGAAGC | CGGTGAGGTA | CTCTTCGCGG | GTGGCGAAAC | CCAGCTCCCC | 4380 |
| GTGGGCACGG | CGGGCCCAGA | AGGAACGCGA | GGTCCCCCAC | CCGGCGAACA | CCCGGCCCCG | 4440 |
| CGCCTTCCGC | CCCCGCTCCC | CGGCGTCGGC | GCTGAGCGCC | GCGGCCAGAC | CGGACAGCAG | 4500 |
| GACCAGGCTG | TGCGGGCTGC | TCACCGGCGC | CCCGCAGATC | GGGGCGATCC | GGCGCACCAT | 4560 |
| CCCCGGATGC | GACACGGCCC | ACTGGTAGGC | GTGGGCCGCG | CCCATCGACC | AGCCCGTGAC | 4620 |
| CAGGGCCAGT | TCCCGTACCC | CCAGCTCCTC | GGTGAGCAGC | CGGTGCTGCG | CCGCGACATT | 4680 |
| GTCTGCGGA | GTGATCAGCG | GAAAGCGGGA | CCCCGACGGG | TGGTTGCCGG | GCGAGCTGGA | 4740 |
| GACCCCGTTG | CCGAAGAGTC | CGGCGGTGAC | GACGCAGTAC | CGCCGGGTGT | CCAGCGGCAG | 4800 |
| CCCCGCACCG | ATCAGCCAGT | CGTACCCGGT | GTGGTCCCGG | CCGAAGAACG | ACGGACAGAG | 4860 |
| CACCACGTTC | GTCCCGTCGG | CGTTCGGCGT | GCCGTACATG | GCGTAACCGA | TCCGGGCGTC | 4920 |
| CCGCAGGACC | TCCCGTCCA | GCAACGGCAG | TTCGTGCATC | TCGAATATGC | GGCATTCAC | 4980 |
| CGCTGACCTC | CTTGTTTCGAT | CCCCCCGGAC | AACAGGTCGG | TCGTGGCCGG | AGACTCAGAG | 5040 |
| CCAGTTGGGG | GCGATCTCGG | TGGCCACACG | CTCCAGGCTG | CGCAGCTGGA | CATCGTCCGG | 5100 |
| GATCAGCCCC | GAGTACTGGC | ACTGGAGCAG | ATACTCCGGA | TCGTGCCGCT | CCACCAGCTT | 5160 |
| CTCGATCATG | CGGTTGATGT | CGTCCGGGGT | GCCGACCCAC | TCCAGCCCCC | GGTCGACCAG | 5220 |
| GGTCTTGTAG | TCCGAGCCGA | TCGGACCCGT | CTCGCCGGTC | GCGCGCAGCG | CCTCGGTGAA | 5280 |
| GCCCCATGGG | CCGAACCAGT | TCTCGAAGAT | GAAGCCGCCG | CCGCGGGACG | CCCAGTGGTG | 5340 |
| GGCCTCGCCG | GAGTCCCGGG | AGACCAGGAC | GTCTTTCATC | ACCCCGACCC | GCTCGCCCCG | 5400 |
| CCGCAGGGTG | CCGTGGCCCC | CCGCCTCGGC | CTCCTCCCGG | TAGATGTCCA | TCAGCCGGGC | 5460 |
| GACGATCTGG | TCGTCCGTGT | TCATCAGGAT | CGGCACCACG | CCCTCCCGGG | CACAGAACCG | 5520 |
| GAACGTGTCC | TCACTGAAGC | TGAACGGCTG | GAAGACGGGC | GGGTGGGGGC | GCTGGTAGGG | 5580 |
| CTTGGGCGCG | ATGCCCCACT | CGCGGATGAC | GCCGTTCCTC | TCGAGGCCCC | GGCCGTAGCG | 5640 |
| GCGCACCGCC | TCGTAGGGGA | ACTCCAGGTC | CGGCACCGGG | ATCGTCCACT | GCTCCCCGGA | 5700 |
| GTGGGTGAAC | GTCTCGGTGC | TCCACGCCTT | CTTGATGATC | TCCCAGTGCT | CCTCGAAGAG | 5760 |
| GGCACGATTG | CGCCGGTCCC | GCTCCCCGGC | GTCGGACAGG | GTGCCGCCGA | CCCCGTACAC | 5820 |
| CTGCCCCATG | ATGTGGGCCC | AGCGCTTCTG | GAACCCGCGC | GCGATCCCCG | CGAAGGCGCG | 5880 |
| GCCCCGGGTC | ATGTGGTCTG | GCATCGCCAG | ATCCTCGGCC | AGCCGCAGCG | GATTGTGCAG | 5940 |
| CGGCAGGACG | TTGGCCATCT | GGCCGACCCG | GATGTGCCGG | GTCTGCATGC | CGAGGTAGAG | 6000 |
| CCCCAGCATG | ATCGGGTTGT | TGGAGACCTC | GAAACCCTCG | GTGTGGAAGT | GGTGCTCGGT | 6060 |

GAAGGACAGT CCCAGTAGC CGAGTTCGTC GGCCGCCTGC GCCTGCCGGG TGAGCTGCCG 6120
GAGCATGTTC TGGTAGTTCT GCGGATTGAC CCCC GCCATA CCCC GCTGGA CCTGCGCATG 6180
ACTGCCGACC GTTGGCAGAT AGAAGAGAAT GGACTTCACC CTGGCTCCTC CCGTTCGCGG 6240
CGCCCTCCAT TGACGTGCGC CGAAAGCGGC TCGACCGTCC CACTCCGCCC TTGAGTTCCG 6300
TCTGACGCCG CGCCAGTCGG CGGGCCGTCC GCCGGGGTGC CCGCCGGGGT CCGCACCCGC 6360
CGGACGGCAC GGCGCGCACC GCGCGCGCGG CGCTTCGGGG CACCGGGGTC GACGGGGTGC 6420
TCAGCGGGAC GTCCAACGGA AGGCAAGCCC CCGTACCCAG CCTGGTCAAG GCGCTCATCG 6480
CCATTCCCTG AGGAGGTCCC GCCTTGACCA CAGCAATCTC CGCGCTCCCG ACCGTGCCCCG 6540
GCTCCGGACT CGAAGCACTG GACCGTGCCA CCCTCATCCA CCCCACCCTC TCCGAAACA 6600
CCGCGGAACG GATCGTGCTG ACCTCGGGGT CCGGCAGCCG GGTCCGCGAC ACCGACGGCC 6660
GGGAGTACCT GGACGCGAGC GCCGTCCTCG GGGTGACCCA GGTGGGGCAC GGCCGGGGCCG 6720
AGCTGGCCCG GGTGCGGGCC GAGCAGATGG CCCGGCTGGA GTACTTCCAC ACCTGGGGGA 6780
CGATCAGCAA CGACCGGGCG GTGGAGCTGG CGGCACGGCT GGTGGGGCTG AGCCCGGAGC 6840
CGCTGACCCG CGTCTACTTC ACCAGCGGCG GGGCCGAGGG CAACGAGATC GCCCTGCGGA 6900
TGGCCCGGCT CTACCACCAC CGGCGCGGGG AGTCCGCCCC TACCTGGATA CTCTCCCGCC 6960
GGTCGGCCTA CCACGGCGTC GGATACGGCA GCGGCGGCGT CACCGGCTTC CCCGCCTACC 7020
ACCAGGGCTT CGGCCCTCC CTCCCGGACG TCGACTTCCT GACCCCGCCG CAGCCCTACC 7080
GCCGGGAGCT GTTCGCCGGT TCCGACGTCA CCGACTTCTG CCTCGCCGAA CTGCGCGAGA 7140
CCATCGACCG GATCGGCCCC GAGCGGATCG CGGCGATGAT CGGCGAGCCG ATC 7193

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 145 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTGACCCGGC CTCCGGGCCT TTCCGCGCAC ACCCAGGGT CCGTGTCGGG GAGTCTGCTG 60
CGCCGGGTGG CGGGCCACTA TCCCACGGG GTGGTCCTGG TCACCGGTCC GGCCGAGGCT 120
CCGGGGCAGC CGCCGCCCCG CATGG 145

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| ATGTCCGTGG | CATCGGCCGG | TATGACGGAC | GAGCAGCGCA | AGGCGGTCAT | CACCGCGTAC | 60 |
| TTCAAGGCGT | TCGACAACGG | CGGCGTCGGC | AGCGACGGCA | CCCCCGCGAT | CGACTACTTC | 120 |
| GCCGAGGACG | CGGTCTTCTT | CTTCCCCAAG | TGGGGTCTGG | CCCGGGGCAA | GTCCGAGATC | 180 |
| GCCCGGCTCT | TCGACGACCT | CGGGGGCACC | ATCCGCTCGA | TCACCCACCA | TCTGTGGTCC | 240 |
| GTCAACTGGA | TTCTGACCGG | GACCGAATC | CTCGCCGCGG | AGGGCACCAC | CCACGGTGAG | 300 |
| CACCGGGACG | GGCCGTGGCG | GGCGGGTGAC | CCCGAGTGGG | CCGCCGGGCG | CTGGTGCACG | 360 |
| GTCTACGAGG | TGCGGGACTT | CCTCGTCCAC | CGGGCCTTCG | TCTATCTGGA | CCCCGATTAC | 420 |
| GCGGGCAAGG | ACACCGCGCG | TTACCCGTGG | CTG | | | 453 |

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1032 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| ATGTCCCGCT | CTCCGCCCGA | GTCCCCGGCC | GGTTCCGTGT | CCGCCGCGGT | TCCGCGTCCG | 60 |
| CCGGTCCGCG | CCCTGCGGGA | CCTTCCGGTC | AGTGCCAGG | GGCTCGGCTG | CCTGCCGACC | 120 |
| ACCGACTTCT | ACGGACGCCC | GGACCGCGCC | CGGGCGACGG | CCACCATCCG | CGCCGCCGTC | 180 |
| GACGCCGGGG | TCACCCTGCT | GGACACCGCC | GACGTCCAGG | GGCTCGGCGC | CGGTGAGGAG | 240 |
| CTGCTCGGAC | GGCGGTCGC | GGGCCGCCGG | GACGAGGTGC | TGATCGCCAC | CAAGTTCGGC | 300 |
| ATGGTGCGCT | CGTCCGACGG | CGCCTCCCAG | GGCTTGTGCG | GCGAGCCGTC | CTACGTCCGC | 360 |
| GCGGCCTGCG | AACGGTCCCT | GCGTCGTCTC | GGCACCAGC | GCATCGACCT | GTACTACGAG | 420 |
| CACTGGACGG | ACCCGGCGGT | GCCGATCGAG | GAGACCGTGG | GTGCGGTGGC | CGAGCTGGTG | 480 |
| CGCGAGGGCA | AGGTCCGCAG | GCTCGGTCTC | TCCGAGCCCT | CCGCGGCCAC | GCTGCGCCGG | 540 |
| GCGGACGCGG | TGACCCCGGT | GACGGCGGTG | CAGAGCGAGT | GGAGCCTGTG | GTGCGCGGGG | 600 |
| ATCGAGGACG | AGGTGGTGCC | CGTCTGCCGG | GAGCTGGGGA | TCGGGATCGT | CGCTTACGCC | 660 |

| | | | | | | |
|------------|------------|------------|-------------|-------------|------------|------|
| CCTCTGGGAC | GGGGTTTTCT | CACCGGCACC | ATCCGCACCA | CCGACGATCT | GGGGGACGAG | 720 |
| GACTTCCGCC | GGGGCCAGCC | CCGGTTCAGC | GCTCCGGCCC | TCGCGCGCAA | CCGCTCGTTG | 780 |
| CTGCACCGGC | TGCGCCCGGT | CGCGGACGGT | CTGGGGCTGA | CCCTGGCACA | GCTCGCGCTC | 840 |
| GCCTGGGTGC | ACCACCGGGG | CGAGGACGTC | GTCCCCGATCC | CGGGCACC GC | GAACCCGGCC | 900 |
| CATCTCGCGG | ACAATCTCGC | CGCCGCCTCG | ATCCGGCTGG | ACGACCGGTC | CCTCGCGGAG | 960 |
| GTGACGGCCG | CGATCTCCCA | CCCGGTGTCC | GGGGAGCGGT | ACACCCCGGC | ATTGCTCGCC | 1020 |
| ATGATCGGCA | AC | | | | | 1032 |

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 984 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

| | | | | | | |
|------------|------------|-------------|-------------|-------------|-------------|-----|
| GTGGAATGCC | GCATATTCGA | GATCGACGAA | CTGCCGTTGC | TGGACGGGGA | GGTCCTGCGG | 60 |
| GACGCCCCGA | TCGGTTACGC | CATGTACGGC | ACGCCGAACG | CCGACGGGAC | GAACGTGGTG | 120 |
| CTCTGTCCGT | CGTTCTTCGG | CCGGGACCAC | ACCGGGTACG | ACTGGCTGAT | CGGTGCGGGG | 180 |
| CTGCCGCTGG | ACACCCGGCG | GTA CTGCGTC | GTCACCGCCG | GACTCTTCGG | CAACGGGGTC | 240 |
| TCCAGCTCGC | CCGGCAACCA | CCCGTCGGGG | TCCCGCTTTC | CGCTGATCAC | TCCGCAGGAC | 300 |
| AATGTGCGCG | CGCAGCACCG | GCTGCTCACC | GAGGAGCTGG | GGGTACGGGA | ACTGGCCCTG | 360 |
| GTCACGGGCT | GGTCGATGGG | CGCGGCCAC | GCCTACCAGT | GGGCCGTGTC | GCATCCGGGG | 420 |
| ATGGTGCGCC | GGATCGCCCC | GATCTGCGGG | GCGCCGGTGA | GCAGCCCGCA | CAGCCTGGTC | 480 |
| CTGCTGTCCG | GTCTGGCCGC | GGCGCTCAGC | GCCGACGCCG | GGGAGCGGGG | GCGGAAGGCG | 540 |
| GCGGGCCGGG | TGTTCGCCGG | GTGGGGGACC | TCGCGTTCCCT | TCTGGGCCCCG | CCGTGCCCCAC | 600 |
| CGGGAGCTGG | GTTTCGCCAC | CCGCGAGGAG | TACCTCACC | GCTTCTGGGA | GCAGGTCTTC | 660 |
| CTCTCCGGGC | CCGGCGCCGC | GGATCTGCTC | ACCATGGTGC | GCACCTGGGA | GAACACGGAT | 720 |
| GTGGGGGCGA | CACCCGGGGC | CGGGGGGAGC | GTCGAGGCGG | CGCTGGCCTC | CGTCACGGCG | 780 |
| CGGGCCGTGG | TGCTGCCGGG | CGCCCTGGAC | GTGTGTTTCG | CCGTCGAGGA | CGAGAAGCGG | 840 |
| GTGGCCGATC | TGCTGCCGTA | TGCCTCGCTG | GAGGTGATCC | CGGGAGTGTG | GGGGCATCTC | 900 |
| GCGGGGTCCG | GGGGGTGCGC | CGCCGACCGG | GAGTTCATCG | GGGGCGCGCT | GCGGCGGCTG | 960 |
| CTGGACAGCC | CGGTGGACGG | GGGC | | | | 984 |

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1182 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

| | |
|---|------|
| GTGAAGTCCA TTCTCTTCTA TCTGCCAACG GTCGGCAGTC ATGCGCAGGT CCAGCGGGGT | 60 |
| ATGGCGGGGG TCAATCCGCA GAACTACCAG AACATGCTCC GGCAGCTCAC CCGGCAGGCG | 120 |
| CAGGCGGGCG ACGAACTCGG CTA CTGCTTCA CCGAGCACCA CTTCCACACC | 180 |
| GAGGGTTTCG AGGTCTCCAA CAACCCGATC ATGCTGGGGC TCTACCTCGG CATGCAGACC | 240 |
| CGGCACATCC GGGTCGGCCA GATGGCCAAC GTCCTGCCGC TGCACAATCC GCTGCGGCTG | 300 |
| GCCGAGGATC TGGCGATGCT CGACCACATG ACCCGGGGCC GCGCCTTCGT CGGGATCGCG | 360 |
| CGCGGGTTCC AGAAGCGCTG GGCCGACATC ATGGGGCAGG TGTACGGGGT CGGCGGCACC | 420 |
| CTGTCCGACG CCGGGGAGCG GGACCGGCGC AATCGTGCCC TCTTCGAGGA GCACTGGGAG | 480 |
| ATCATCAAGA AGGCGTGAC GACCGAGACG TTCACCCACT CCGGGGAGCA GTGACGATC | 540 |
| CCGGTGCCCG ACCTGGAGTT CCCCTACGAG GCGGTGCGCC GCTACGGCCG GGGCCTCGAC | 600 |
| GAGAACGGCG TCATCCGCGA GGTGGGCATC GCGCCCAAGC CCTACCAGCG CCCCCACCCG | 660 |
| CCCGTCTTCC AGCCGTTTCA CTTTCTGAG GACACGTTCC GGTTCTGTGC CCGGGAGGGC | 720 |
| GTGGTGCCGA TCCTGATGAA CACCGACGAC CAGATCGTCG CCCGGCTGAT GGACATCTAC | 780 |
| CGGGAGGAGG CCGAGGCGGC GGGCCACGGC ACCCTGCGGC GGGGCGAGCG GGTCGGGGTG | 840 |
| ATGAAGGACG TCCTGGTCTC CCGGGACTCC GGCGAGGCC ACCACTGGGC GTCCCGCGGC | 900 |
| GGCGGCTTCA TCTTCGAGAA CTGGTTCGGC CCCATGGGCT TCACCGAGGC GCTGCGCGCG | 960 |
| ACCGGCGAGA CCGGTCCGAT CGGCTCGGAC TACAAGACCC TGGTCGACCG GGGGCTGGAG | 1020 |
| TGGGTCGGCA CCGCGACGA CATCAACCGC ATGATCGAGA AGCTGGTGGA GCGGCACGAT | 1080 |
| CCGAGTATC TGCTCCAGTG CCAGTACTCC GGGCTGATCC CGCACGATGT CCAGCTGCGC | 1140 |
| AGCCTGGAGC TGTGGGCCAC CGAGATCGCC CCCAACTGGC TC | 1182 |

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 660 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

| | |
|---|-----|
| GTGCCCCGGCT CCGGACTCGA AGCACTGGAC CGTGCCACCC TCATCCACCC CACCCTCTCC | 60 |
| GGAAACACCG CGGAACGGAT CGTGCTGACC TCGGGGTCCG GCAGCCGGGT CCGCGACACC | 120 |
| GACGGCCGGG AGTACCTGGA CGCGAGCGCC GTCCTCGGGG TGACCCAGGT GGGCCACGGC | 180 |
| CGGGCCGAGC TGGCCCGGGT CGCGGCCGAG CAGATGGCCC GGCTGGAGTA CTTCCACACC | 240 |
| TGGGGGACGA TCAGCAACGA CCGGGCGGTG GAGCTGGCGG CACGGCTGGT GGGGCTGAGC | 300 |
| CCGGAGCCGC TGACCCGCGT CTA CTACTTCACC AGCGGCGGGG CCGAGGGCAA CGAGATCGCC | 360 |
| CTGCGGATGG CCCGGCTCTA CCACCACCGG CGCGGGGAGT CCGCCCGTAC CTGGATACTC | 420 |
| TCCCCCGGGT CGGCCTACCA CGGCGTCGGA TACGGCAGCG GCGGCGTCAC CGGCTTCCCC | 480 |
| GCCTACCACC AGGGCTTCGG CCCCTCCCTC CCGGACGTCG ACTTCCTGAC CCCGCCGCAG | 540 |
| CCCTACCGCC GGGAGCTGTT CGCCGGTTCC GACGTCACCG ACTTCTGCCT CGCCGAAGTG | 600 |
| CGCGAGACCA TCGACCGGAT CGGCCCGGAG CGGATCGCGG CGATGATCGG CGAGCCGATC | 660 |

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

| | |
|-----------------------------|----|
| CTGACGCTGC AGGAGGAAGT CCCGC | 25 |
|-----------------------------|----|

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGGGGCGAGG ACGTCGTCCC GATCC

25

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAGCCCCTGG ACGTCGGCGG TGTCC

25

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GACGGTGCAT GCTCAGCAGG GAGCG

25

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 972 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ATGACCTCAG TGGACTGCAC CGCGTACGGC CCCGAGCTGC GCGCGCTCGC CGCCCGGCTG 60
CCCCGGACCC CCCGGGCCGA CCTGTACGCC TTCCTGGACG CCGCGCACAC AGCCGCCGCC 120
TCGCTCCCCG GCGCCCTCGC CACCGCGCTG GACACCTTCA ACGCCGAGGG CAGCGAGGAC 180
GGCCATCTGC TGCTGCGCGG CCTCCCGGTG GAGGCCGACG CCGACCTCCC CACCACCCCG 240
AGCAGCACCC CGGCGCCCGA GGACCGCTCC CTGCTGACCA TGGAGGCCAT GCTCGGACTG 300
GTGGGCCGCC GGCTCGGTCT GCACACGGGG TACCGGGAGC TGCCTCGGG CACGGTCTAC 360
CACGACGTGT ACCCGTCGCC CGGCGCGCAC CACCTGTCCT CGGAGACCTC CGAGACGCTG 420
CTGGAGTTCC ACACGGAGAT GGCTTACCAC CGGCTCCAGC CGAACTACGT CATGCTGGCC 480
TGCTCCCGGG CCGACCACGA GCGCACGGCG GCCACACTCG TCGCCTCGGT CCGCAAGGCG 540
CTGCCCCCTG TGGACGAGAG GACCCGGGCC CGGCTCCTCG ACCGGAGGAT GCCCTGCTGC 600
GTGGATGTGG CCTTCCGCGG CGGGGTGGAC GACCCGGGCG CCATCGCCCA GGTCAAACCG 660
CTCTACGGGG ACGCGGACGA TCCCTTCCTC GGGTACGACC GCGAGCTGCT GGCGCCGGAG 720
GACCCCGCGG ACAAGGAGGC CGTCGCCGCC CTGTCCAAGG CGCTCGACGA GGTCAACGGAG 780
GCGGTGTATC TGGAGCCCGG CGATCTGCTG ATCGTCGACA ACTTCCGCAC CACGCACGCG 840
CGGACGCCGT TCTCGCCCCG CTGGGACGGG AAGGACCGCT GGCTGCACCG CGTCTACATC 900
CGCACCAGAC GCAATGGACA GCTCTCCGGC GGCGAGCGCG CGGGCGACGT CGTCGCCTTC 960
ACACCGCGCG GC 972

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Thr Arg Pro Pro Gly Leu Ser Ala His Thr His Gly Ser Val Ser
1 5 10 15
Gly Ser Leu Leu Arg Arg Val Ala Gly His Tyr Pro Thr Gly Val Val
20 25 30
Leu Val Thr Gly Pro Ala Glu Ala Pro Gly Gln Pro Pro Pro Ala Met
35 40 45

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 151 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ser Val Ala Ser Ala Gly Met Thr Asp Glu Gln Arg Lys Ala Val
1 5 10 15
Ile Thr Ala Tyr Phe Lys Ala Phe Asp Asn Gly Gly Val Gly Ser Asp
 20 25 30
Gly Thr Pro Ala Ile Asp Tyr Phe Ala Glu Asp Ala Val Phe Phe Phe
 35 40 45
Pro Lys Trp Gly Leu Ala Arg Gly Lys Ser Glu Ile Ala Arg Leu Phe
50 55 60
Asp Asp Leu Gly Gly Thr Ile Arg Ser Ile Thr His His Leu Trp Ser
65 70 75 80
Val Asn Trp Ile Leu Thr Gly Thr Glu Leu Leu Ala Ala Glu Gly Thr
 85 90 95
Thr His Gly Glu His Arg Asp Gly Pro Trp Arg Ala Gly Asp Pro Glu
 100 105 110
Trp Ala Ala Gly Arg Trp Cys Thr Val Tyr Glu Val Arg Asp Phe Leu
 115 120 125
Val His Arg Ala Phe Val Tyr Leu Asp Pro Asp Tyr Ala Gly Lys Asp
130 135 140
Thr Ala Arg Tyr Pro Trp Leu
145 150

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 344 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ser | Arg | Ser | Pro | Pro | Glu | Ser | Pro | Ala | Gly | Ser | Val | Ser | Ala | Ala |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Val | Pro | Arg | Pro | Pro | Val | Arg | Ala | Leu | Arg | Asp | Leu | Pro | Val | Ser | Ala |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gln | Gly | Leu | Gly | Cys | Leu | Pro | Thr | Thr | Asp | Phe | Tyr | Gly | Arg | Pro | Asp |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Arg | Ala | Arg | Ala | Thr | Ala | Thr | Ile | Arg | Ala | Ala | Val | Asp | Ala | Gly | Val |
| | | 50 | | | | 55 | | | | | 60 | | | | |
| Thr | Leu | Leu | Asp | Thr | Ala | Asp | Val | Gln | Gly | Leu | Gly | Ala | Gly | Glu | Glu |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Leu | Leu | Gly | Arg | Ala | Val | Ala | Gly | Arg | Arg | Asp | Glu | Val | Leu | Ile | Ala |
| | | | | 85 | | | | | 90 | | | | 95 | | |
| Thr | Lys | Phe | Gly | Met | Val | Arg | Ser | Ser | Asp | Gly | Ala | Ser | Gln | Gly | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Cys | Gly | Glu | Pro | Ser | Tyr | Val | Arg | Ala | Ala | Cys | Glu | Arg | Ser | Leu | Arg |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Arg | Leu | Gly | Thr | Asp | Arg | Ile | Asp | Leu | Tyr | Tyr | Gln | His | Trp | Thr | Asp |
| | | 130 | | | | 135 | | | | | 140 | | | | |
| Pro | Ala | Val | Pro | Ile | Glu | Glu | Thr | Val | Gly | Ala | Val | Ala | Glu | Leu | Val |
| 145 | | | | 150 | | | | | | 155 | | | | 160 | |
| Arg | Glu | Gly | Lys | Val | Arg | Arg | Leu | Gly | Leu | Ser | Glu | Pro | Ser | Ala | Ala |
| | | | 165 | | | | | 170 | | | | | 175 | | |
| Thr | Leu | Arg | Arg | Ala | Asp | Ala | Val | His | Pro | Val | Thr | Ala | Val | Gln | Ser |
| | | 180 | | | | | 185 | | | | | | 190 | | |
| Glu | Trp | Ser | Leu | Trp | Ser | Arg | Gly | Ile | Glu | Asp | Glu | Val | Val | Pro | Val |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Cys | Arg | Glu | Leu | Gly | Ile | Gly | Ile | Val | Ala | Tyr | Ala | Pro | Leu | Gly | Arg |
| | | 210 | | | | 215 | | | | | 220 | | | | |
| Gly | Phe | Leu | Thr | Gly | Thr | Ile | Arg | Thr | Thr | Asp | Asp | Leu | Gly | Asp | Glu |
| 225 | | | | 230 | | | | | | 235 | | | | 240 | |
| Asp | Phe | Arg | Arg | Gly | Gln | Pro | Arg | Phe | Ser | Ala | Pro | Ala | Leu | Ala | Arg |
| | | | 245 | | | | | 250 | | | | | 255 | | |
| Asn | Arg | Ser | Leu | Leu | His | Arg | Leu | Arg | Pro | Val | Ala | Asp | Gly | Leu | Gly |
| | | 260 | | | | | | 265 | | | | | 270 | | |
| Leu | Thr | Leu | Ala | Gln | Leu | Ala | Leu | Ala | Trp | Leu | His | His | Arg | Gly | Glu |

275 280 285
Asp Val Val Pro Ile Pro Gly Thr Ala Asn Pro Ala His Leu Ala Asp
290 295 300
Asn Leu Ala Ala Ala Ser Ile Arg Leu Asp Asp Arg Ser Leu Ala Glu
305 310 315 320
Val Thr Ala Ala Ile Ser His Pro Val Ser Gly Glu Arg Tyr Thr Pro
325 330 335
Ala Leu Leu Ala Met Ile Gly Asn
340

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 328 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Glu Cys Arg Ile Phe Glu Ile Asp Glu Leu Pro Leu Leu Asp Gly
1 5 10 15
Glu Val Leu Arg Asp Ala Arg Ile Gly Tyr Ala Met Tyr Gly Thr Pro
20 25 30
Asn Ala Asp Gly Thr Asn Val Val Leu Cys Pro Ser Phe Phe Gly Arg
35 40 45
Asp His Thr Gly Tyr Asp Trp Leu Ile Gly Ala Gly Leu Pro Leu Asp
50 55 60
Thr Arg Arg Tyr Cys Val Val Thr Ala Gly Leu Phe Gly Asn Gly Val
65 70 75 80
Ser Ser Ser Pro Gly Asn His Pro Ser Gly Ser Arg Phe Pro Leu Ile
85 90 95
Thr Pro Gln Asp Asn Val Ala Ala Gln His Arg Leu Leu Thr Glu Glu
100 105 110
Leu Gly Val Arg Glu Leu Ala Leu Val Thr Gly Trp Ser Met Gly Ala
115 120 125
Ala His Ala Tyr Gln Trp Ala Val Ser His Pro Gly Met Val Arg Arg
130 135 140

Ile Ala Pro Ile Cys Gly Ala Pro Val Ser Ser Pro His Ser Leu Val
145 150 155 160
Leu Leu Ser Gly Leu Ala Ala Ala Leu Ser Ala Asp Ala Gly Glu Arg
165 170 175
Gly Arg Lys Ala Ala Gly Arg Val Phe Ala Gly Trp Gly Thr Ser Arg
180 185 190
Ser Phe Trp Ala Arg Arg Ala His Arg Glu Leu Gly Phe Ala Thr Arg
195 200 205
Glu Glu Tyr Leu Thr Gly Phe Trp Glu Gln Val Phe Leu Ser Gly Pro
210 215 220
Gly Ala Ala Asp Leu Leu Thr Met Val Arg Thr Trp Glu Asn Thr Asp
225 230 235 240
Val Gly Ala Thr Pro Gly Ala Gly Gly Ser Val Glu Ala Ala Leu Ala
245 250 255
Ser Val Thr Ala Arg Ala Val Val Leu Pro Gly Ala Leu Asp Val Cys
260 265 270
Phe Ala Val Glu Asp Glu Lys Arg Val Ala Asp Leu Leu Pro Tyr Ala
275 280 285
Ser Leu Glu Val Ile Pro Gly Val Trp Gly His Leu Ala Gly Ser Gly
290 295 300
Gly Ser Ala Ala Asp Arg Glu Phe Ile Gly Gly Ala Leu Arg Arg Leu
305 310 315 320
Leu Asp Ser Pro Val Asp Gly Gly
325

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 394 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Lys Ser Ile Leu Phe Tyr Leu Pro Thr Val Gly Ser His Ala Gln
1 5 10 15
Val Gln Arg Gly Met Ala Gly Val Asn Pro Gln Asn Tyr Gln Asn Met

20 25 30
Leu Arg Gln Leu Thr Arg Gln Ala Gln Ala Ala Asp Glu Leu Gly Tyr
35 40 45
Trp Gly Leu Ser Phe Thr Glu His His Phe His Thr Glu Gly Phe Glu
50 55 60
Val Ser Asn Asn Pro Ile Met Leu Gly Leu Tyr Leu Gly Met Gln Thr
65 70 75 80
Arg His Ile Arg Val Gly Gln Met Ala Asn Val Leu Pro Leu His Asn
85 90 95
Pro Leu Arg Leu Ala Glu Asp Leu Ala Met Leu Asp His Met Thr Arg
100 105 110
Gly Arg Ala Phe Val Gly Ile Ala Arg Gly Phe Gln Lys Arg Trp Ala
115 120 125
Asp Ile Met Gly Gln Val Tyr Gly Val Gly Gly Thr Leu Ser Asp Ala
130 135 140
Gly Glu Arg Asp Arg Arg Asn Arg Ala Leu Phe Glu Glu His Trp Glu
145 150 155 160
Ile Ile Lys Lys Ala Trp Thr Thr Glu Thr Phe Thr His Ser Gly Glu
165 170 175
Gln Trp Thr Ile Pro Val Pro Asp Leu Glu Phe Pro Tyr Glu Ala Val
180 185 190
Arg Arg Tyr Gly Arg Gly Leu Asp Glu Asn Gly Val Ile Arg Glu Val
195 200 205
Gly Ile Ala Pro Lys Pro Tyr Gln Arg Pro His Pro Pro Val Phe Gln
210 215 220
Pro Phe Ser Phe Ser Glu Asp Thr Phe Arg Phe Cys Ala Arg Glu Gly
225 230 235 240
Val Val Pro Ile Leu Met Asn Thr Asp Asp Gln Ile Val Ala Arg Leu
245 250 255
Met Asp Ile Tyr Arg Glu Glu Ala Glu Ala Ala Gly His Gly Thr Leu
260 265 270
Arg Arg Gly Glu Arg Val Gly Val Met Lys Asp Val Leu Val Ser Arg
275 280 285
Asp Ser Gly Glu Ala His His Trp Ala Ser Arg Gly Gly Gly Phe Ile
290 295 300
Phe Glu Asn Trp Phe Gly Pro Met Gly Phe Thr Glu Ala Leu Arg Ala
305 310 315 320
Thr Gly Glu Thr Gly Pro Ile Gly Ser Asp Tyr Lys Thr Leu Val Asp
325 330 335
Arg Gly Leu Glu Trp Val Gly Thr Pro Asp Asp Ile Asn Arg Met Ile

340 345 350
Glu Lys Leu Val Glu Arg His Asp Pro Glu Tyr Leu Leu Gln Cys Gln
355 360 365
Tyr Ser Gly Leu Ile Pro His Asp Val Gln Leu Arg Ser Leu Glu Leu
370 375 380
Trp Ala Thr Glu Ile Ala Pro Asn Trp Leu
385 390

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 220 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Pro Gly Ser Gly Leu Glu Ala Leu Asp Arg Ala Thr Leu Ile His
1 5 10 15
Pro Thr Leu Ser Gly Asn Thr Ala Glu Arg Ile Val Leu Thr Ser Gly
20 25 30
Ser Gly Ser Arg Val Arg Asp Thr Asp Gly Arg Glu Tyr Leu Asp Ala
35 40 45
Ser Ala Val Leu Gly Val Thr Gln Val Gly His Gly Arg Ala Glu Leu
50 55 60
Ala Arg Val Ala Ala Glu Gln Met Ala Arg Leu Glu Tyr Phe His Thr
65 70 75 80
Trp Gly Thr Ile Ser Asn Asp Arg Ala Val Glu Leu Ala Ala Arg Leu
85 90 95
Val Gly Leu Ser Pro Glu Pro Leu Thr Arg Val Tyr Phe Thr Ser Gly
100 105 110
Gly Ala Glu Gly Asn Glu Ile Ala Leu Arg Met Ala Arg Leu Tyr His
115 120 125
His Arg Arg Gly Glu Ser Ala Arg Thr Trp Ile Leu Ser Arg Arg Ser
130 135 140
Ala Tyr His Gly Val Gly Tyr Gly Ser Gly Gly Val Thr Gly Phe Pro
145 150 155 160

Ala Tyr His Gln Gly Phe Gly Pro Ser Leu Pro Asp Val Asp Phe Leu
165 170 175
Thr Pro Pro Gln Pro Tyr Arg Arg Glu Leu Phe Ala Gly Ser Asp Val
180 185 190
Thr Asp Phe Cys Leu Ala Glu Leu Arg Glu Thr Ile Asp Arg Ile Gly
195 200 205
Pro Glu Arg Ile Ala Ala Met Ile Gly Glu Pro Ile
210 215 220

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 324 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Thr Ser Val Asp Cys Thr Ala Tyr Gly Pro Glu Leu Arg Ala Leu
1 5 10 15
Ala Ala Arg Leu Pro Arg Thr Pro Arg Ala Asp Leu Tyr Ala Phe Leu
20 25 30
Asp Ala Ala His Thr Ala Ala Ala Ser Leu Pro Gly Ala Leu Ala Thr
35 40 45
Ala Leu Asp Thr Phe Asn Ala Glu Gly Ser Glu Asp Gly His Leu Leu
50 55 60
Leu Arg Gly Leu Pro Val Glu Ala Asp Ala Asp Leu Pro Thr Thr Pro
65 70 75 80
Ser Ser Thr Pro Ala Pro Glu Asp Arg Ser Leu Leu Thr Met Glu Ala
85 90 95
Met Leu Gly Leu Val Gly Arg Arg Leu Gly Leu His Thr Gly Tyr Arg
100 105 110
Glu Leu Arg Ser Gly Thr Val Tyr His Asp Val Tyr Pro Ser Pro Gly
115 120 125
Ala His His Leu Ser Ser Glu Thr Ser Glu Thr Leu Leu Glu Phe His
130 135 140
Thr Glu Met Ala Tyr His Arg Leu Gln Pro Asn Tyr Val Met Leu Ala

145 150 155 160
Cys Ser Arg Ala Asp His Glu Arg Thr Ala Ala Thr Leu Val Ala Ser
165 170 175
Val Arg Lys Ala Leu Pro Leu Leu Asp Glu Arg Thr Arg Ala Arg Leu
180 185 190
Leu Asp Arg Arg Met Pro Cys Cys Val Asp Val Ala Phe Arg Gly Gly
195 200 205
Val Asp Asp Pro Gly Ala Ile Ala Gln Val Lys Pro Leu Tyr Gly Asp
210 215 220
Ala Asp Asp Pro Phe Leu Gly Tyr Asp Arg Glu Leu Leu Ala Pro Glu
225 230 235 240
Asp Pro Ala Asp Lys Glu Ala Val Ala Ala Leu Ser Lys Ala Leu Asp
245 250 255
Glu Val Thr Glu Ala Val Tyr Leu Glu Pro Gly Asp Leu Leu Ile Val
260 265 270
Asp Asn Phe Arg Thr Thr His Ala Arg Thr Pro Phe Ser Pro Arg Trp
275 280 285
Asp Gly Lys Asp Arg Trp Leu His Arg Val Tyr Ile Arg Thr Asp Arg
290 295 300
Asn Gly Gln Leu Ser Gly Gly Glu Arg Ala Gly Asp Val Val Ala Phe
305 310 315 320
Thr Pro Arg Gly